UNIVERSITY OF MARYLAND

National Transportation Center

Project ID: NTC2015-MU-R-1

UNDERSTANDING REGIONAL DISPARITIES IN PUBLIC TRANSIT PERFORMANCE USING REALTIME TRANSIT DATA

Final Report

by

Alex Karner
Georgia Institute of Technology

Aaron Golub
Portland State University

Celeste Chavis
Morgan State University

for

National Transportation Center at Maryland (NTC@Maryland)
1124 Glenn Martin Hall
University of Maryland
College Park, MD 20742

March 2016

ACKNOWLEDGEMENTS

This project was funded by the National Transportation Center @ Maryland (NTC@Maryland), one of the five National Centers that were selected in a nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT). Dominick Tribone and David Barker at the Metropolitan Boston Transportation Authority provided invaluable assistance with the data and methods employed here. We are also grateful to Shiva Narooienezhad for her help with data assembly.

DISCLAIMER

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The Moving Ahead for Progress in the 21st Century (MAP-21) act, signed by President Obama in 2012, indicated that explicit and objective performance targets should be set in the areas of congestion reduction, system reliability, and environmental sustainability, among others. The science of performance assessment is being driven forward by emerging data sources and declining storage costs. For public transit systems in particular, advances in the spatial and temporal representation of public transit supply, combined with emerging datasets that provide highly resolved information about the location and socio-economic nature of demand are opening up analytical possibilities for operations, planning, and optimization. These newer datasets allow for performance measures to be calculated at the level of individual stops in a transit network at any moment of the day. The key sources that reveal these precise understandings of public transit systems are route, schedule, and realtime position and trip updates from the general transit feed specification (GTFS) and GTFS-realtime formats. The former provides a static picture of expected service, while the latter provides precise information about vehicle positions and expected delays system-wide.

The goal of this study is to demonstrate how these realtime data can be applied to provide insight into the distribution of public transit service across a region. A common approach to measuring transit performance is to perform a "gap analysis" that examines mismatches between supply and demand, with the goal of identifying locations where demand for transit is high and supply is low. Prior to this work, these measures were often created using *expected* or scheduled service which provides no information about the service actually delivered to patrons. For this study, we assembled both GTFS and GTFS realtime feeds for the metropolitan Boston area, gleaned from the Massachusetts Bay Transportation Authority (MBTA), to provide a more realistic picture of how transit performs in a region by incorporating real-time delay information into the analysis.

For comparison, a traditional gap analysis was conducted using expected service information obtained through the static GTFS feed for bus trips during the morning peak period (7-9am). Using American Community Survey data, demand was operationalized using standardized race and income proportions that are summed and grouped into deciles at the level of census tract. A frequency-based measure of public transit supply was also deployed at the census tract level; supply was measured as the number of trips per hour reachable within 0.5 miles of a transit stop as measured by distance along the pedestrian network. Consistent with the literature, transit supply and demand were found to be correlated. However, the traditional gap analysis does not take into account actual service provision. A novel measure of supply was developed by incorporating bus stop level delay using GTFS-realtime data, aggregated by census tract, where supply is measured as the difference of the standardized supply score from the standardized delay score. The results of the gap analysis with delay incorporated show a much more even distribution of transit service across demand deciles exists and that there is in fact no significant difference across demand deciles. This shows that when service delivery is considered, the results of schedule-based gap analysis is reversed due to the diminished quality of service cased by actual delays.

1.0 INTRODUCTION

The Moving Ahead for Progress in the 21st Century (MAP-21) act, signed by President Obama in 2012, indicated that explicit and objective performance targets should be set in the areas of congestion reduction, system reliability, and environmental sustainability, among others. New data sources are now available to better understand the dynamic nature of public transit performance which can assist in understanding how to improve such systems. This project contributes to federal goals by leveraging these emerging data sources for improved performance evaluation, increasing the likelihood that transit supply and demand will be well-matched, and a more equitable transportation system will be achieved.

Recent advances in the spatial and temporal representation of public transit supply combined with emerging datasets that provide highly-resolved information about the location and socio-economic nature of public transit demand are opening up analytical possibilities for public transit operations, planning, accessibility, and optimization (Catala, Downing, & Hayward, 2011; Porter, Kim, & Ghanbartehrani, 2014). Whereas prior datasets only allowed for coarse spatial and temporal analysis on the scale of a census tract or transportation analysis zone (TAZ) over a period of several hours, these newer datasets allow for accessibility and other performance characteristics to be calculated at the level of individual stops or routes in a transit network at a particular instant on a particular day. Overall performance measures can still be derived through appropriate aggregations of the disaggregate measures (Karner, 2016).

The key resource facilitating these precise understandings of public transit systems come from individual transit operators making data available in the general transit feed specification (GTFS). The GTFS was created to standardize the reporting and updating of public transit route and schedule information and forms the basis for transit directions found on Google Transit (Catala et al., 2011). Most transit providers maintain a static GTFS feed that provides a precise accounting of expected operational transit service. Some operators combine this with a real-time GTFS feed that provides detailed data on the movements of individual transit vehicles. Combining both of these datasets where possible offers a powerful way to understand how actual public transit service compares to planned or expected service, revealing how delay or service reliability changes through time, space and along a transit route. This project will leverage these new data tools in an evaluation of public transportation service equity – it will add a metric of delay to understand how delay varies across a route and how that affects different demographic groups differently

This report will begin with an overview of the existing research, and how this project advances this area. It will then describe the data and methods used to develop our metrics of transit supply, delay, and transit need. We finish with results and a discussion of our analysis.

1.1 BACKGROUND AND LITERATURE REVIEW

Transportation systems create the opportunity to access important destinations beyond our immediate surroundings. In large metropolitan areas where uses are dispersed widely in space, a lack of transportation can mean a lack of opportunities for communities and individuals (Blumenberg & Waller, 2003; Ihlanfeldt & Sjoquist, 1998; Sanchez & Brenman, 2008). Therefore, there is important concern for the distribution of transportation investments and services. Even though automobiles offer better mobility performance in most areas of the United States and are thus important for many households (Taylor & Paul, 1995) public transportation services offer households more affordable mobility options compared to car ownership, as well as provide essential services for the substantial numbers of people who cannot drive (Public Policy Institute of California, 2004). Achieving adequate and equitable public transit services thus has been an ongoing challenge in the US for decades (Bullard et al., 2004; Pucher, 1982; Rosenbloom & Altshuler, 1977). The quality of public transportation services vary greatly by geography and demographics and level of service disparities can drive legal challenges to transit agencies; bus scheduling and vehicle overcrowding were central to the Los Angeles Bus Riders' Union successful complaint against the Los Angeles Metropolitan Transportation Authority (Mann, 2004). Thus, significant attention is being paid to measuring and analyzing such disparities across a range of dimensions.

The work presented here sits at the intersection of three large bodies of research. The first includes a vast array of performance measures for public transit systems, specifically accessibility and mobility measures and measures of transit supply at different scales in time, space, and along routes and at bus stops. The second focuses on understanding and measuring disparities in transportation service and the access it creates. The third is working to add real-time information to understand delay and reliability to transit performance measures and understand how delay and reliability affect other aspects of system performance and demand for transit.

The first body of literature incorporates a wide range of methods and measures to understand the nature of transit performance, accessibility and mobility and how these data can be visualized, used as ingredients to the public transit planning process and to improve and optimize transit systems investments and services. A few representative examples from a larger set of literature include: Benenson et al. (2010), Bhat et al. (2005), Lei and Church (2010), Lei and Goulias (2012), Martin et al. (2008), Mavoa et al. (2012), Mesbah et al. (2012), Rood (1999), Mamun and Lownes (2011a), Tribby and Zandbergen (2012), and Wu and Hine (2003). Some of this literature includes static measure of transit availability (e.g. Rood (1999)), while others add time-of-day measures incorporating actual service schedules, congestion and other dynamic issues (e.g. Martin et al. (2008) and Bhat et al. (2005), among others.). This work is important in establishing a baseline understanding of different transit supply measures which can then inform further measures of disparities.

Research looking at disparities in transit service is quite expansive, and reflects a range of methods, time and spatial scales and levels of detail of service provision. A few representative examples from a larger set of literature include: Blair et al. (2013), Bocarejo et al. (2012), Currie (2010), Farber et al. (2014), Farber et al. (2016), Ferguson et al. (2012), Foth et al. (2013), Fransen et al. (2015), Golub and Martens (2014), Grengs (2010, 2013, 2015), Grengs and Levine (2010), Kaplan et al. (2014), Karner (2016), Kawabata (2009), Kawabata and Shen (2007), Mamun and Lownes (2011b), Minocha et al. (2012), Murray and Davis (2001), Thomopoulos and Grant-Muller (2013), and Wells and Thill (2012). Many of these include accessibility measures reflecting travel time from an origin to other parts of the network (e.g. Farber et al. (2014), Karner (2016) or Grengs (2010)), while some measure a gap between transit need in an origin and the supply or performance at that origin (e.g. Currie (2010), Mamun and Lownes (2011b)).

Finally, recent work has added a variability measure to traditional understandings of transit performance, recognizing that reliability and delay can be a strong detriment to potential ridership even when added service is adequate. With the emergence of systematized transit schedule information through GTFS and its incorporation into GIS, researchers are better able to assess disparities in service by time of day and in space (A. Karner, 2016), while the additional real-time vehicle location data allows better resolution of delay along a route and at stops (Chen, Yu, Zhang, & Guo, 2009). Much of this work deals with understanding how delay and reliability affect transit operations, while others seek to measure how delays affects the overall benefit of transit investments. A few representative examples from a larger set of literature on these issues include: Barabino et al (2015), Chakrabarti (2015), Chang (2010), Chen et al. (2009), Fonzone et al. (2015), Nagatani (2012), Peña and Moreno (2014), Saberi et al. (2013), and Tirachini (2013, 2014).

This project will incorporate detailed measures of delay into a measure of transit supply within an equity analysis framework, comparing supply and demand and effectively combining these three areas of research. Our work will identify areas in a city or region where actual public transit service is performing poorly relative to expected service (through a measure of real-time delay) and will provide a method for other cities to conduct similar analyses. We will also determine whether there is systematic variation in space in terms of this delay effect conditioned upon demographic characteristics across the city or region. The results can provide information about where and why public transit services differ from expectation and will provide public transit planners with valuable information about the equity of any departures from expectation and will also highlight areas of the system that require improvement and attention.

2.0 DATA AND METHODS

Understanding mismatches between the location of public transit demand and the location and the timing of public transit supply has long been a concern of the literature. This study explores the mismatches in supply and demand for the Massachusetts Bay Transit Authority (MBTA), the public operator in the greater Boston, Massachusetts area. The transit network consists of commuter rail, subway, bus (local and express) and boats (ferries and commuter boats). MBTA serves over 4.8 million people and covers a service area over 3200 square miles. Typical weekday unlinked ridership is around 1.3 million trips - resulting in nearly 390 million trips annually. MBTA maintains 9 subway routes, 174 bus routes, 14 commuter rail routes and 3 ferry routes (Massachusetts Bay Transportation Authority, 2014). Moreover, MBTA provides consistent real-time data in GTFS format for all equipped modes as well as alert data. Each MBTA bus is equipped with GPS. Through a partnership with NextBus, MBTA makes available bus arrival predictions at each bus stop. GTFS-realtime provides alerts, vehicle location and arrival prediction data. However for the data be meaningful, it must be extrapolated using static-GTFS which includes schedule and trip planning data. MBTA and IBI (2014) provides more information on MBTA's GTFS-realtime API. These data sources, their preparation, and the analyses undertaken, are described below.

To operationalize public transit supply and demand, we gathered data from the American Community Survey (ACS) five-year estimates along with public transit route and schedule information gleaned from GTFS and GTFS-realtime feed maintained by MBTA. Software was written to integrate GTFS-realtime and static feeds within a database and to visualize differences between them.

2.1 GTFS DATABASE DEVELOPMENT

Two open source tools, GTFS Database (GTFSDB) and GTFS-realtime to Database (GTFSrDB), available on GitHub were adapted to collect and store static and real-time GTFS data for MBTA into an SQLite database. GTFSDB[1] and GTFSrDB[2] contain Python code that loads GTFS data into a SQLite relational database. GTFSrDB supports all three types of GTFS-realtime feeds: Trip Updates, Service Alerts, and Vehicle Positions. Trip Updates represents fluctuations in the time table by providing predicted arrival information for stops along the route. Service Alerts provides updates whenever there is a disruption on the network that need to be communicated to public. Lastly, Vehicle Positions provides vehicle locations every 30 seconds. The static GTFS database is May 27, 2015 – September 4, 2015. GTFS-realtime database was collected from July 1 to August 13, 2015.

2.2 TRANSIT SUPPLY

Our measure of public transit supply is fundamentally frequency-based and will ultimately be calculated at the level of a census tract so that it can be compared with demographics. Geographic areas should score more highly on our measure if they have more transit vehicles running within specific time periods. For this initial analysis, we assessed service during the morning peak period (7-9am). To quantify transit supply and generate a measure of trips per hour during this period, we used the "BetterBusBuffers" (BBB) tool created and maintained by ESRI and implemented within ArcMap.[3] In brief, the tool employs the static GTFS feed to count the number of arrivals/departures reachable within pedestrian service areas constructed around transit stops and stations. The output from a BBB run includes detailed arrival rates (trips per hour) for each service area during a defined time period.

Service areas represent a dramatic improvement over standard Euclidian buffers. Whereas Euclidian buffers simply deploy a circle around a point of interest (e.g. a transit stop) to quantify locations accessible to it, service areas represent locations reachable from a given point along the actual pedestrian network. In other words, service areas take into account network distances whereas Euclidian buffers do not. It is very common for service areas to overlap, especially in dense urban areas where transit stops are located in close proximity. BBB calculates separate frequency measures for all intersected and flattened service areas, to avoid double counting. The intersection creates many service area fragments that maintain their association with particular stops. The BBB frequency value for any fragment includes all unique trips visiting stops associated with that service area within the time period of interest. If a single trip visits multiple stops associated with that same service area, it is only counted once. Thus, the tool ensures that trips are not double-counted. However, since each stop or station is associated with multiple different

[1] For more information see https://github.com/OpenTransitTools/gtfsdb
[2] For more information see https://github.com/mattwigway/gtfsrdb
[3] More information on the tool and its use can be found here: http://www.transit.melindamorang.com/overview_BetterBusBuffers.html .

service areas, a single trip contributes to multiple service area counts. Using service areas also allows individual bus stops to contribute to apparent service in multiple tracts as opposed to just the tract in which it is physically located.

In order to execute the BBB tool, we had to create a reasonable pedestrian network for use in ArcMap with which to create service areas. The network should include a representation of streets and other physical infrastructure (e.g., sidewalks, pedestrian-only paths) used by pedestrians to access the transit network. One low-cost possibility is to use OpenStreetMap (OSM) data for the area of interest.[4] While quite advanced in Europe, using these data for the United States is somewhat problematic. Most of the US data were imported from the US Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) format and have to be manually cleaned by members of the public engaged in the OpenStreetMap community.[5] In many parts of the US, this cleaning simply has not occurred, or has occurred only sporadically. Despite these limitations, OSM is the best source of free data and can be used as long as these limitations are kept in mind and the data are checked for obvious flaws.

OSM data for specific geographic areas, as opposed to the entire world, are routinely exported and stored in various formats in several locations on the web. OSM data at the state level are available.[6] We imported OSM data for entire state of Massachusetts, as MBTA service extends beyond the borders of the city of Boston. All street links within 0.5 miles of a transit junction were subsequently selected to form the basis for the pedestrian network. We then prepared a subset of the data containing streets accessible to pedestrians by creating a new feature class containing only elements that pedestrians can use.[7] Standard ArcMap Network Analyst tools can then be used to create a suitable network. We ran the BBB tool to generate trips per hour estimates within 0.5 miles of all bus stops. These were subsequently aggregated to the census tract level by first converting all service areas to points and taking the mean over all service areas within a tract.

2.2.1 Calculating bus-stop-level delay

A key goal of this study was to quantify the value added associated with moving to real-time as opposed to static measures of public transit supply. Using the collected GTFS-realtime data, we estimated a measure of delay at the bus stop level using the trip update feed that was subsequently aggregated to the census tract to match the frequency based supply measure described above. For the purposes of this study we focused solely on bus stop-level delay estimates.[8] The focus on bus

---

[4] Alternative, and comparatively more costly, options include StreetMap Premium sold by ESRI. See: http://www.esri.com/data/streetmap .

[5] The OpenStreetMap wiki describes the numerous problems with direct TIGER imports. See: http://wiki.openstreetmap.org/wiki/TIGER_fixup .

[6] http://download.geofabrik.de/north-america.html

[7] Tags in the attribute table of the feature class indicate the functional class to which a particular street or road belongs (residential, highway, etc.). Features whose tags indicated functional classes not traversable by pedestrians were eliminated. Other polyline features represented utility rights-of-way, streams, and rivers. These were eliminated as well.

[8] Heavy rail delays will be addressed in future work. Additionally, MBTA operates several other modes including ferry, streetcar, and commuter rail, but the level of service characteristics of these modes make them less relevant to transit dependent populations.

is necessitated by the manner in which trip updates are returned in the GTFS-realtime feed. During the period over which data were collected, each bus trip update was associated with at most one stop time update whereas each heavy rail trip update included all stops remaining in the trip. For bus, the stop time update provided was for the next stop along that trip's route. For the heavy rail trip updates, each trip update query returns a stop time update for *each* stop that remains for that particular trip.[9] In either case, because API queries were made every 30 seconds, it is possible for the next stop or station to receive multiple updates related to expected delay as the transit vehicle moves between these terminal locations. Converting these different temporal measures of delay into a single representative value for each stop and station is a challenge.

The approach we took to calculate delay at the bus stop level recognizes that for a single trip,[10] multiple updates can be provided for a single stop as a vehicle approaches, encounters delay, makes up time, and ultimately arrives. Arguably what matters most for the transit user is whether the bus arrives in a particular location on schedule. For this reason we chose to use the most accurate estimate of delay residing in the database, which we defined as the expected delay for the trip update provided when the bus is as close to the next stop as possible. In other words, it is the final trip update provided before the bus passes that stop and begins announcing delay for the subsequent stops. The process of delay estimation is illustrated in Figure 1. As a bus departs stop 333 its expected delay in arriving at stop 334 is zero. It encounters some unexpected congestion along the way, increasing its delay estimate to 240 seconds. Subsequently, it is able to make up some time and manages to reduce its delay to 60 seconds. In the final trip update announced before arriving at stop 334 (highlighted in red), the arrival delay is 60 seconds.[11] Although a higher delay had been announced earlier in the trip, the final delay announcement will be closest to reality and that is what is used here. Note that the resolution of delay in this GTFS-realtime feed was whole minutes. These correspond to expected arrivals and departures from the static feed, which occur only on whole minutes and zero seconds.

[9] As of early 2016, this approach (providing a stop time update for every remaining stop) is now MBTA's standard practice for *all modes*. The GTFS-realtime standard generally builds in more flexibility in data reporting than does the static GTFS standard so that what is reported can change over time. This structure is helpful for GTFS-realtime consumers engaged in application development (since app users positioned at later stops can be provided with an immediate update that does not require intermediate calculations), but is less useful for the present study. In general, the more recent data structure dramatically increases storage requirements.

[10] In GTFS parlance a trip refers to the journey of a single transit vehicle from its origin stop or station to its terminal stop or station.

[11] Arrival delays are only precise to the minute. Although on its face this seems like a limitation, the GTFS static schedule embodies minute-level precision as well.

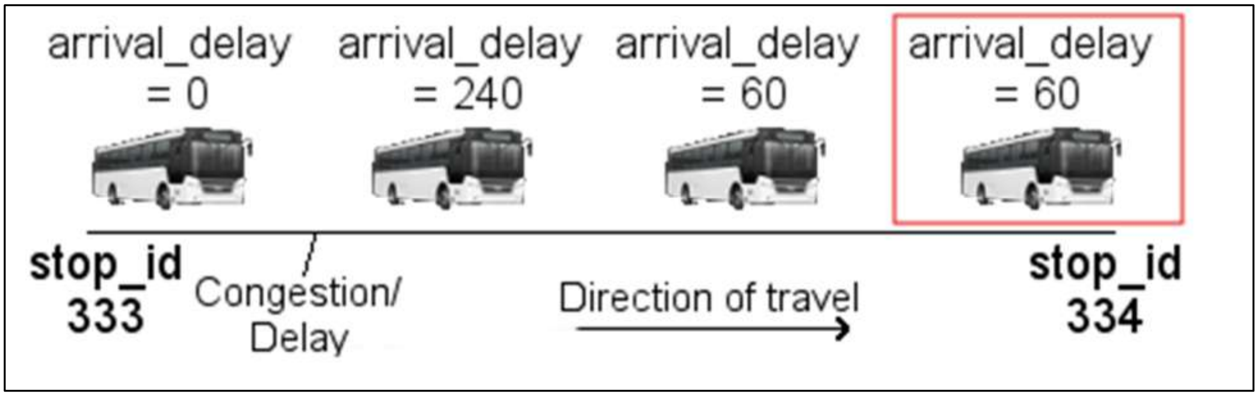

Figure 1. Illustration of variation in announced delay as a transit vehicle (bus) moves between two stops. Trip updates were retrieved every 30 seconds over the analysis period. The arrival delay highlighted in red would be the delay associated with this particular date-trip-stop combination.

To generate this delay from the GTFS-realtime database, we needed to query the database to return only the last announced delay for each unique combination of trip start date, trip, and stop. To accomplish this task, we took advantage of the fact that subsequent trip updates were numbered sequentially with a unique identifier, meaning that higher values of the identifier occur later in time. For every unique date-trip-stop combination, we extracted the trip update identifier and then joined that back to the stop time updates to extract the arrival delay announced at the last stop. From an initial population of 49.4 million stop time updates, we extracted 8.11 million unique combinations as described above. Again, the total population of stop time updates exceeds what is needed from an analysis perspective because multiple stop time updates can be announced for an individual stop as a transit vehicle traverses a route and encounters changing conditions. The approach we used here to estimate delay is more theoretically satisfying than taking a straight mean of all possible delays, because vehicles experiencing delay will be more likely to announce subsequent updates for a single stop. If we used every announced delay, we would effectively be weighting those stops that experience delay more heavily in our calculations. Additionally, initial delays can be recovered over the course of a trip.

The result of these queries is that we extract an estimate of stop-level delay for a single trip across multiple days. These delays can sensibly be combined to generate measures of central tendency. We estimated means, medians, ranges, standard deviations, and coefficients of variation to characterize the variability and reliability of transit service at the stop level. We also performed quality control on the delay data to ensure that any missing stops were sensible. According to the corresponding static GTFS feed, there are 7,874 bus stops in the MBTA network. Of these, 7,517 received at least one stop time update from the real-time feed over the period of data collection, 377 received no update, and 20 received an update but did not have a corresponding entry in the static feed. The 20 stops receiving an update were likely temporary stops entered into the real-time feed, which can take precedence over the static feed. The 377 stops which did not receive an update could represent historical stops no longer in use but which are still included in the feed.

2.3 TRANSIT DEMAND

Consistent with practice in the literature, we defined transit demand based on the demographic characteristics of census tracts. Data were gleaned from the American Community Survey (ACS) 2010-2014 five-year estimates.[12] Although there are obvious limitations with this approach, namely that the particular demographics chosen are arbitrary and different results may be obtained if different demographics are employed, employing this new data source for demonstration purposes is appropriate. We chose race and income as the primary demographics of interest. Specifically, we calculated the proportion of tract residents that are people of color (i.e. total population minus the non-Hispanic white alone population), and the proportion of households in a tract with incomes below 80% of the area median income. The reference geography for the income category was the Boston-Cambridge-Newton, MA metropolitan statistical area. According to the 2010-2014 ACS estimates, the median income in the MSA was $74,670; 80% of that value is $59,736.

Although household income is clearly an important determinant of transit demand to the extent that it determines automobile holdings and required trip distances and frequencies (Pucher & Renne, 2003), race is a less obvious predictor of demand per se. Much prior research has demonstrated an independent effect of race on travel behavior even when controlling for more obvious predictors (Giuliano, 2005; Kockelman, 1997; Liu, 2000; Mauch & Taylor, 1997). This residual effect might be due to the combination of segregated metropolitan areas that tend to place people of color into particular places in regions and the likelihood that these places will be well-served by transit. There are at least two additional reasons to include race as an analytical category. First, there is a long tradition in the analysis of public services distribution of that highlights disparate treatment of people of color (blacks in particular) (Lineberry, 1977; Wilson, 1990). Additionally, federal civil rights law prohibits discrimination on the basis of race, color, or national origin in the provision of transit service (Alex Karner & Golub, 2015). Understanding how the distribution of transit service varies across a region with respect to race and income simultaneously should provide insight regarding demand and equity.

[12] The US Census Bureau offers major demographic data pre-joined with census geographic information for multiple ACS, Decennial Census, and County Business Patterns years. Data are available here: https://www.census.gov/geo/maps-data/data/tiger-data.html

3.0 RESULTS

3.1 BUS DELAY ESTIMATES

Figure 2 summarizes the 8.11 million estimates of stop-level delay as gleaned from MBTA's real-time feed over the entire study period. The data summarized here are for illustration purposes only; they differ from those used in our final delay calculations because they include multiple estimates for individual stops, depending on the exact dates that were included in the data collection. The mode of the delay is clearly zero, with a long tail extending in the direction of increasing delay as well as a smaller proportion of arrivals that are ahead of schedule. Although early arrivals can also be a problem for transit passengers, if the bus departs *before* the indicated departure time, for subsequent analyses we bottom-coded all of the delay values that were less than zero to zero under the assumption that the vehicle would hold at the stop until the scheduled departure time.

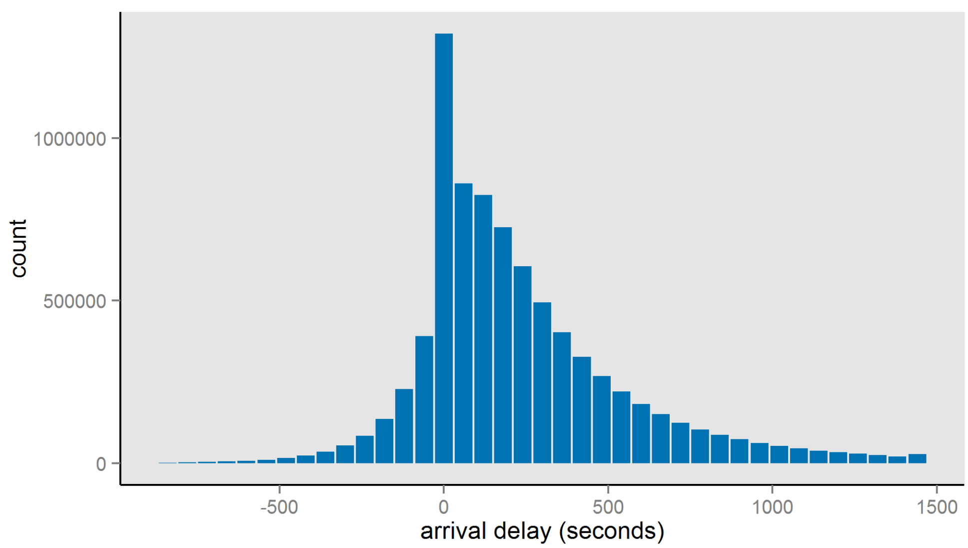

Figure 2. Distribution of measured stop-level delay for buses in the MBTA system over the entire study period.

It is sensible to summarize all of the delay information about a particular stop at the level of individual stops to identify measures of central tendency and variation. Collapsing the full set of stop-level estimates for bus delay results in a single estimate for each stop, even if multiple routes serve particular stops. These results are shown in Figure 3 for a number of measures of central tendency and dispersion. Here, the mean delay (panel A) is no longer constrained to a resolution of minutes, since we are taking an average over a large number of observations. In general, taking this average reduces the maximum observation substantially, but moves the overall mean to the right. The average stop in the system experienced a mean delay of 227 seconds, or about four minutes, over the entire time period of the observations. Results for median arrival delay, standard deviation, and coefficient of variation are also shown. The coefficient of variation in particular, shows a right skew, indicating a small group of stops with relatively high values. This result could indicate a problem with service reliability in these locations.

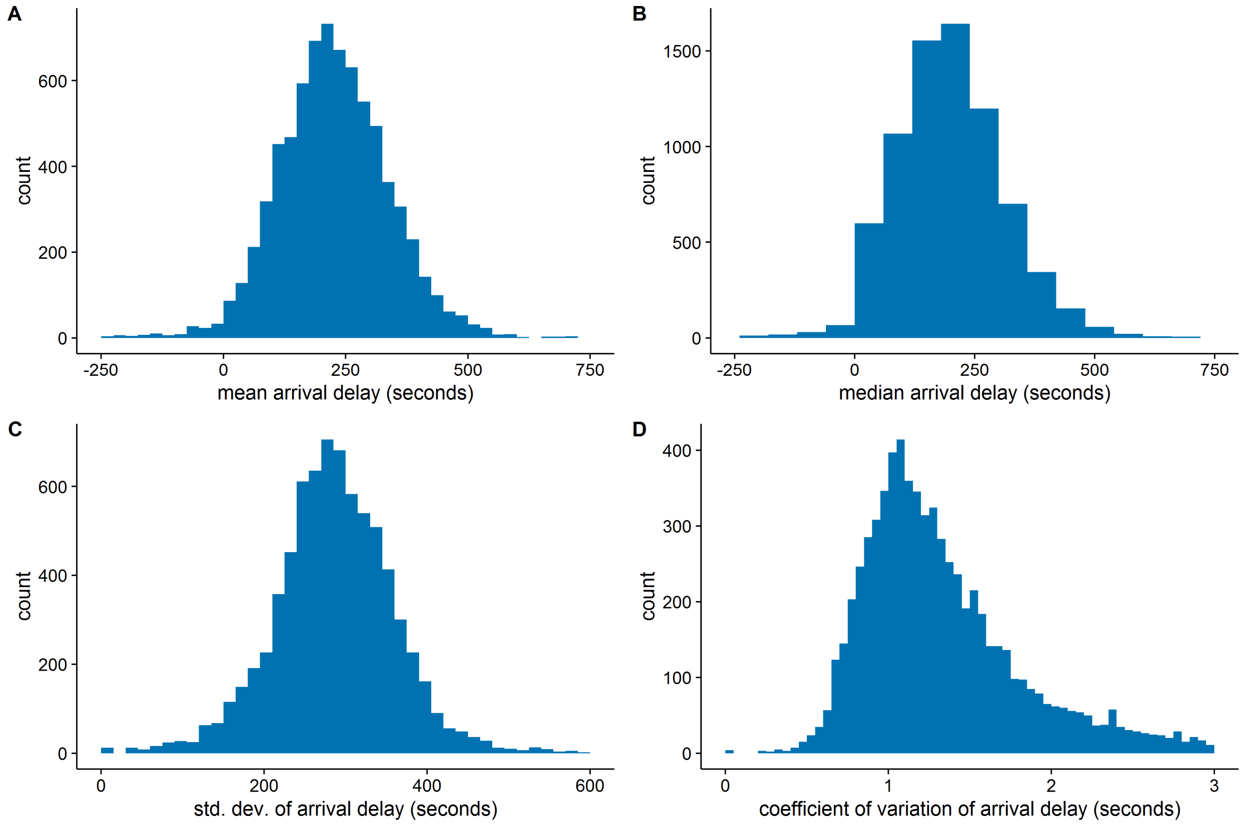

Figure 3. Distribution of measured stop-level delay for buses in the MBTA system aggregated to the stop level. Each bus stop receives a single estimate for (a) mean delay, (b) median delay, (c) standard deviation of delay, and (d) coefficient of variation of arrival delay based on its performance over the entire study period.

3.2 GAP ANALYSIS

A public transit gap analysis typically seeks to identify locations where the demand for transit is high, as operationalized using demographic characteristics, and the supply is low. Identifying whether these gaps exist at the regional level can help to focus the attention of a transit agency on gap closure efforts. However, as noted above, these studies often find that places with apparently high presumed demand are also well-served by transit. The measures used are often relative, so that high-demand locations have the best transit service in the region on the metric used to operationalize service. But the results tell us very little about whether that service is *adequately* serving the needs of transit dependent populations. Further, almost all existing gap analysis studies rely on expected or scheduled service, rather than that actually delivered. The innovation embodied in this study is to use real-time delay information to adjust expected service and provide a more realistic picture of how transit actually functions in a region.

Figure 4 illustrates the results of the gap analysis using expected service (mean trips per hour, calculated by taking the mean frequency results calculated for 3/4 mile service areas around bus stops at the census tract level). Again, demand is operationalized for the purposes of this study using standardized race and income proportions that are summed and grouped into deciles for those tracts that receive some level of transit service. The results confirm expectations. In general, transit supply and demand increase together, as judged by the increasing median values of transit service. These results reflect the historic (although shifting) concentration of people of color and low-income people in central city areas across the United States. The measures employed here are simply transit supply measures, they do not account for the nature of accessibility provided by these trips. It is possible that service frequency is high, but required destinations are not reachable. Additionally, demand was calculated here based on proportions of low-income people and people of color in a census tract. As such, the demand deciles do not account for the absolute counts of these populations within tracts. It may be the case that proportions are high in areas with concentrated transit service, but counts are low. These situations would be better captured with population-weighed metrics.

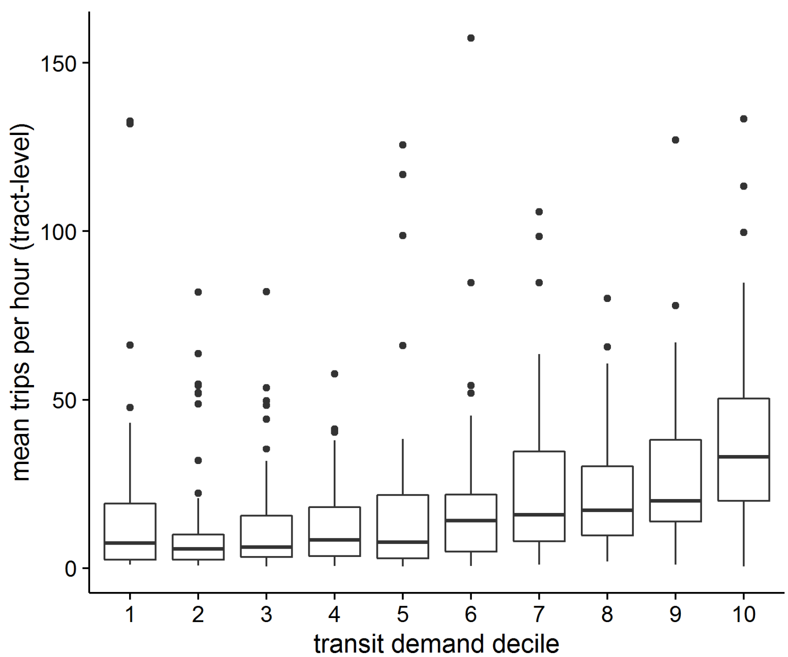

Figure 4. Gap analysis examining the relationship between mean trips per hour in census tracts with transit service and the corresponding transit demand decile. This analysis does not account for delay and focuses only on expected or scheduled service. The deciles measure transit need by combining unweighted, standardized income and racial demographics into a single score. The results are presented as a boxplot. The solid horizontal bar in the box represents the median value for a particular demand decile. The top and bottom of the box represent the 75th and 25th percentile values of transit supply, respectively. The vertical lines emanating from the boxes capture values that fall within 1.5 times of the interquartile range (i.e. the range between the 25th and 75th percentile values). Outliers beyond this range are plotted as points.

We also performed a one-way analysis of variance (ANOVA) and pairwise t-tests to determine whether the apparent differences displayed in Figure 4 are statistically significant. The omnibus test revealed statistically significant different mean values of transit supply between transit demand deciles ($F_{9,574} = 8.00$, $p < 0.001$). Table 1 shows the pairwise t-test results. Each cell is testing the null hypothesis that the group means are equal. The results indicate that deciles seven through 10 differ from many (and in the case of decile 10, all) of the lower demand deciles. In general, supply means across the lower demand deciles are not significantly different from each other.

Table 1. Results of pairwise t-tests (*p*-values shown) comparing mean values of transit supply across transit demand deciles.

transit demand decile	1	2	3	4	5	6	7	8	9
2	NS								
3	NS	NS							
4	NS	NS	NS						
5	NS	NS	NS	NS					
6	NS	NS	NS	NS	NS				
7	<0.05	<0.01	<0.01	<0.01	NS	NS			
8	NS	<0.05	<0.05	<0.05	NS	NS	NS		
9	<0.05	<0.001	<0.001	0.00072	<0.05	NS	NS	NS	
10	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.01

Note: NS = Not significant.

Now that the standard findings from the gap analysis literature have been confirmed, the study's main hypothesis can be tested: do the findings from the gap analysis change when delay is taken into account? Given the history of disparities in service provision by demographics in the US and elsewhere, we would expect that, when delay is taken into account, conditions would appear worse for transit dependent populations than they would seem based on the expected schedule. Figure 5 provides a visual summary of the delay metric calculated for bus service for this study by mapping the results over the MBTA bus service area. Each plotted point is a single bus stop whose delay information has been aggregated over the entire period of study. The stops were subsequently grouped into quintiles. Again, for each trip servicing that stop, only the delay announcement made when the bus was nearest to the stop was considered in the calculation so that stops experiencing heavy delays would not receive additional updates. Additionally, the plotted location of the bus stop does not correspond to its physical location. Because many bus stops are physically proximate, it was necessary to jitter or shift the points so that stops which would otherwise overlap others could be visualized appropriately.

No clear spatial patterns emerge from the mapping exercise. It appears as though most of the stops that experience low delay are located in more peripheral locations of the service area. But there are also low delay stops scattered within high-service locations in Boston and Cambridge. Similarly, there are stops reporting high delays located in the more distant locations of Waltham and Quincy. Although the map represents each bus stop in the exact same manner, they actually differ in important ways in that some are much more heavily used than others. It may be the case that stops seeing heavier use, potentially served by multiple routes, are more likely to experience high delays.

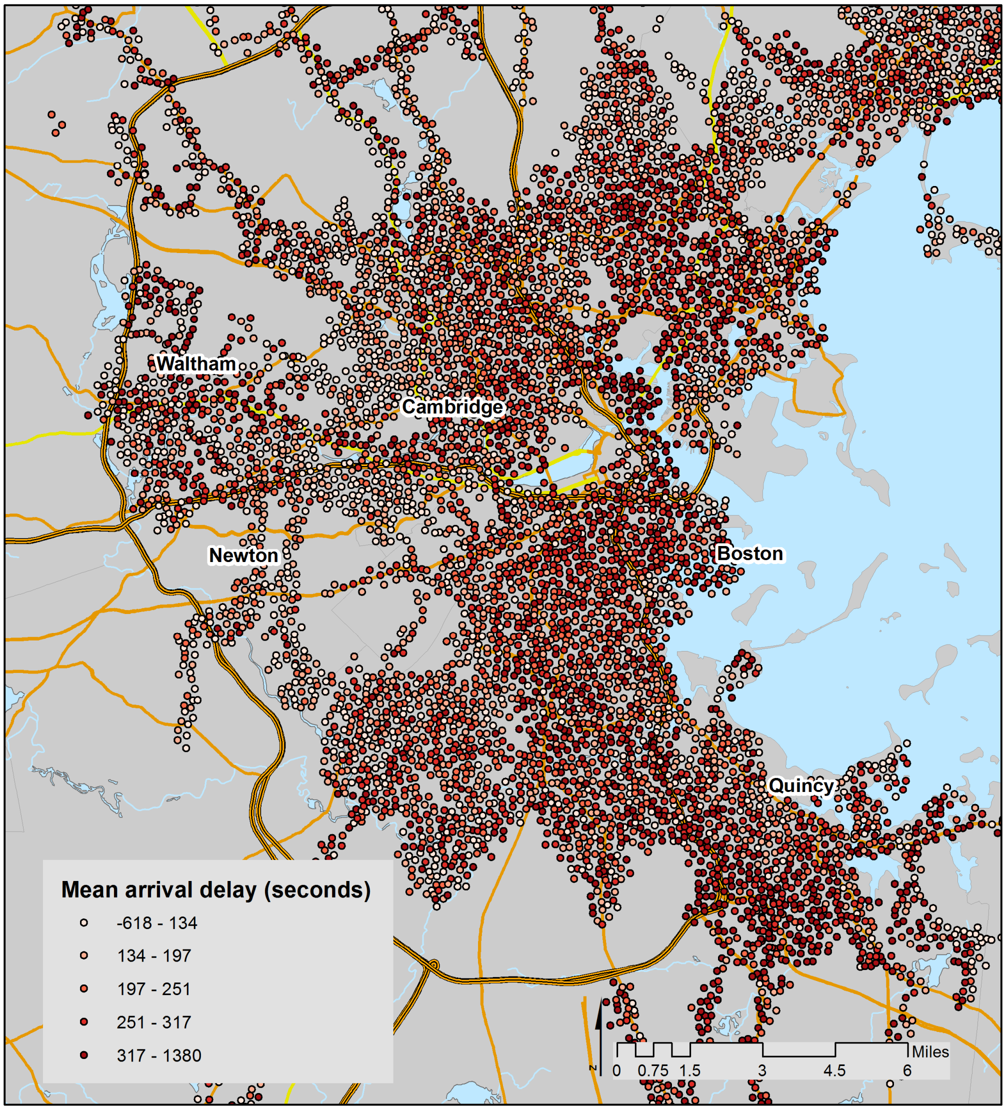

Figure 5. Mean arrival delay by stop over entire dataset. Stops are grouped into quintiles of delay. Note that the location of stops on the map has been scattered to improve visibility; their position does not correspond precisely to their physical location.

To partially guard against this possibility, we incorporated delay information into the gap analysis by aggregating expected delay at the stop level up to the census tract level by taking the mean over all stops. This approach should result in a representative delay value for a tract by averaging over many stops with differing service profiles and frequencies. At the tract level, mean delay and mean trips per hour were subsequently standardized so that they could be compared. The "supply" score for the gap analysis considering delay was calculated by subtracting the standardized delay score from the supply score. High delay values would receive high standardized scores and low delay values would receive the opposite. The effect of delay, then, would be to diminish the apparent performance of particular stop locations if positive, and enhance their performance if negative.

The results for the gap analysis considering delay are summarized in Figure 6. To the best of our knowledge, these results and this figure are the first in the literature to consider actual delay or actual service delivery in the conceptualization of gaps in transit service. In contrast to Figure 4 and much of the existing literature on transit gap analysis, these results show a much more even distribution of transit service across deciles. The omnibus ANOVA test reveals no significant difference between group means ($F_{9,490} = 1.55$, $p = 0.127$). The lowest demand decile receives the highest median standardized supply score (0.225) once delay is taken into consideration. Ranked in order, deciles receiving the highest to lowest levels of service are 1, 3, 10, 6, 8, 9, 5, 7, 4, and 2. Although the results are based on a single month in a particular US region, they clearly demonstrate that consideration of actual service delivery can reverse the findings generated by a typical public transit gap analysis. Although aggregate service may be concentrated in locations that have apparently high transit demands, these locations are also likely to experience relatively high delays, diminishing the quality of service provided and calling into question the utility of relying strictly on schedule-based measures to quantify transit supply.

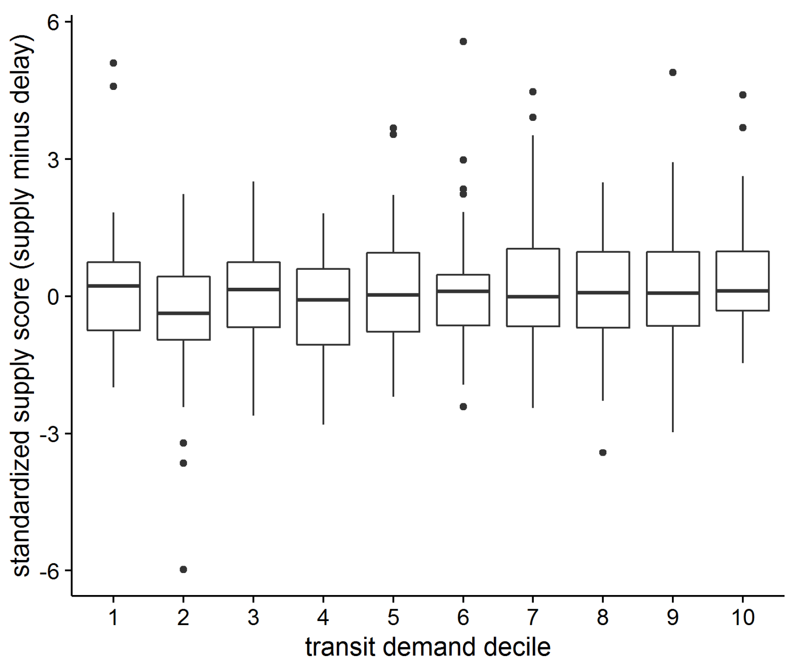

Figure 6. Gap analysis examining the relationship between standardized trips per hour minus standardized delay in census tracts with transit service and the corresponding transit demand decile. This analysis accounts for expected delay. The deciles measure transit need by combining unweighted, standardized income and racial demographics into a single score. The results are presented as a boxplot. The solid horizontal bar in the box represents the median value for a particular demand decile. The top and bottom of the box represent the 75th and 25th percentile values of transit supply, respectively. The vertical lines emanating from the boxes capture values that fall within 1.5 times of the interquartile range (i.e. the range between the 25th and 75th percentile values). Outliers beyond this range are plotted as points.

4.0 CONCLUSIONS

Most public transit planning exercises are conducted using ad hoc processes based on idealized datasets that do not well-represent actual transit service. This study shows that when delay is considered in the transit supply measurement of a gap analysis, the established trend of more supply being provided in areas of high demand vanishes. For transit agencies, this indicates that more frequent service does not necessarily improve the supply of transit to those who need it most. Agencies must also consider schedule adherence and delay.

Though this study was limited in scope it has shed critical insights to the importance of delay and quality of service provisions. Future work may incorporate other independent variables including income, race/ethnicity, educational attainment, vehicle ownership, congestion levels. In addition, temporal considerations can be included to explore time-of-day and seasonal affects in transit provision.

The methods developed may be applied to any agency maintaining GTFS-realtime. Future studies may provide information about where and why public transit services differ from expectation and provide public transit planners with valuable information about the equity of any departures from expectation and highlight areas of the system that require improvement and attention. In addition, if any operational changes are implemented, the methodology can be used to evaluate their performance and impacts.

5.0 REFERENCES

BARABINO, B., DI FRANCESCO, M., & MOZZONI, S. (2015). Rethinking bus punctuality by integrating Automatic Vehicle Location data and passenger patterns. *Transportation Research Part A: Policy and Practice*. http://doi.org/10.1016/j.tra.2015.03.012

BENENSON, I., MARTENS, K., ROFÉ, Y., & KWARTLER, A. (2010). Public transport versus private car GIS-based estimation of accessibility applied to the Tel Aviv metropolitan area. *The Annals of Regional Science*, *47*(3), 499–515. http://doi.org/10.1007/s00168-010-0392-6

BHAT, C. R., GUO, J. Y., SEN, S., & WESTON, L. (2005). *Measuring Access to Public Transportation Services: Review of Customer-Oriented Transit Performance Measures and Methods of Transit Submarket Identification (FHWA/TX08/0-5178-1)*. Center for Transportation Research, The University of Texas at Austin.

BLAIR, N., HINE, J., & BUKHARI, S. M. A. (2013). Analysing the impact of network change on transport disadvantage: A GIS-based case study of Belfast. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2013.06.015

BLUMENBERG, E., & WALLER, M. (2003). The Long Journey to Work : for Working Families: A Federal Transportation Policy for Working Families. *The Brookings Institution Transportation Reform Series*, (July), 1–20.

BOCAREJO S., J. P., & OVIEDO H., D. R. (2012). Transport accessibility and social inequities: a tool for identification of mobility needs and evaluation of transport investments. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2011.12.004

BULLARD, R.; JOHNSON, G.; TORRES, A. (Ed.). (2004). *Highway Robbery: Transportation Racism and New Routes to Equity*. South End Press.

CATALA, M., DOWNING, S., & HAYWARD, D. (2011). *Expanding the Google Transit Feed Specificatoin to Support Operations and Planning*. National Center for Transit Research.

CHAKRABARTI, S. (2015). The demand for reliable transit service: New evidence using stop level data from the Los Angeles Metro bus system. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2015.09.006

CHANG, J. S. (2010). Assessing travel time reliability in transport appraisal. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2009.06.012

CHEN, X., YU, L., ZHANG, Y., & GUO, J. (2009). Analyzing urban bus service reliability at the stop, route, and network levels. *Transportation Research Part A: Policy and Practice*. http://doi.org/10.1016/j.tra.2009.07.006

CURRIE, G. (2010). Quantifying spatial gaps in public transport supply based on social needs. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2008.12.002

FARBER, S., MORANG, M. Z., & WIDENER, M. J. (2014). Temporal variability in transit-based accessibility to supermarkets. *Applied Geography*. http://doi.org/10.1016/j.apgeog.2014.06.012

FARBER, S., RITTER, B., & FU, L. (2016). Space–time mismatch between transit service and observed travel patterns in the Wasatch Front, Utah: A social equity perspective. *Travel Behaviour and Society*. http://doi.org/10.1016/j.tbs.2016.01.001

FERGUSON, E. M., DUTHIE, J., UNNIKRISHNAN, A., & WALLER, S. T. (2012). Incorporating equity into the transit frequency-setting problem. *Transportation Research Part A: Policy and Practice*. http://doi.org/10.1016/j.tra.2011.06.002

FONZONE, A., SCHMÖCKER, J. D., & LIU, R. (2015). A model of bus bunching under reliability-based passenger arrival patterns. *Transportation Research Part C: Emerging Technologies*. http://doi.org/10.1016/j.trc.2015.05.020

FOTH, N., MANAUGH, K., & EL-GENEIDY, A. M. (2013). Towards equitable transit: Examining transit accessibility and social need in Toronto, Canada, 1996-2006. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2012.12.008

FRANSEN, K., NEUTENS, T., FARBER, S., DE MAEYER, P., DERUYTER, G., & WITLOX, F. (2015). Identifying public transport gaps using time-dependent accessibility levels. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2015.09.008

GIULIANO, G. (2005). Low Income, Public Transit, and Mobility. *Transportation Research Record*, *1927*, 63–70. http://doi.org/10.3141/1927-08

GOLUB, A., & MARTENS, K. (2014). Using principles of justice to assess the modal equity of regional transportation plans. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2014.07.014

GRENGS, J. (2010). Job accessibility and the modal mismatch in Detroit. *Journal of Transport Geography*, *18*(1), 42–54. http://doi.org/10.1016/j.jtrangeo.2009.01.012

GRENGS, J. (2013). Nonwork Accessibility as a Social Equity Indicator. *International Journal of Sustainable Transportation*. http://doi.org/10.1080/15568318.2012.719582

GRENGS, J. (2015). *Advancing Social Equity Analysis in Transportation with the Concept of Accessibility - Population Studies Center Research Report 15-848*. Population Studies Center Research Report 15-848.

GRENGS, J., & LEVINE, J. (2010). Intermetropolitan Comparison of Transportation Accessibility: Sorting Out Mobility and Proximity in San Francisco and Washington, D.C.

Journal of Planning Education and Research, *29*(4), 427–443. http://doi.org/10.1177/0739456X10363278

IHLANFELDT, K. R., & SJOQUIST, D. L. (1998). The spatial mismatch hypothesis : a review of recent studies and their implications for welfare reform. *Housing Policy Debate.*, *9*(4), 849–892.

KAPLAN, S., POPOKS, D., PRATO, C. G., & CEDER, A. (2014). Using connectivity for measuring equity in transit provision. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2014.04.016

KARNER, A. (2016). *Assessing public transit service equity using route-level accessibility measures and public data* - Working Paper, Georgia Tech School of City & Regional Planning.

KARNER, A., & GOLUB, A. (2015). Comparison of Two Common Approaches to Public Transit Service Equity Evaluation. *Transportation Research Record: Journal of the Transportation Research Board*, *2531*, 170–179. http://doi.org/10.3141/2531-20

KAWABATA, M. (2009). Spatiotemporal dimensions of modal accessibility disparity in Boston and San Francisco. *Environment and Planning A*, *41*(1), 183–198. http://doi.org/10.1068/a4068

KAWABATA, M., & SHEN, Q. (2007). Commuting Inequality between Cars and Public Transit: The Case of the San Francisco Bay Area, 1990-2000. *Urban Studies*, *44*(9), 1759–1780. http://doi.org/10.1080/00420980701426616

KOCKELMAN, K. (1997). Travel Behavior as Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from San Francisco Bay Area. *Transportation Research Record: Journal of the Transportation Research Board*, *1607*, 116–125. http://doi.org/10.3141/1607-16

LEI, T. L., & CHURCH, R. L. (2010). Mapping transit-based access: integrating GIS, routes and schedules. *International Journal of Geographical Information Science*, *24*(1), 283–304. http://doi.org/10.1080/13658810902835404

LEI, T. L., & GOULIAS, K. G. (2012). Opportunity-Based Dynamic Transit Accessibility in Southern California. *Transportation Research Record Journal of the Transportation Research Board Transportation Research Board of the National Academies*, (2276), 26–37. http://doi.org/10.3141/2276-04

LINEBERRY, R. L. (1977). *Equality and Urban Policy: The Distribution of Municipal Public Services*. SAGE Publications, Inc.

LIU, F. (2000). *Environmental Justice Analysis: Theories, Methods, and Practice*. Boca Raton, FL: Lewis Publishers.

MAMUN, A., & LOWNES, N. (2011a). Measuring Service Gaps. *Transportation Research Record: Journal of the Transportation Research Board*, *2217*, 153–161. http://doi.org/10.3141/2217-19

MAMUN, A., & LOWNES, N. (2011b). A Composite Index of Public Transit Accessibility. *Journal of Public Transportation*, *14*(2).

MANN, E. (2004). Los Angeles bus riders derail the MTA. In *Highway Robbery: Transportation Racism and New Routes to Equity*.

MARTIN, D. J., JORDAN, H., & RODERICK, P. (2008). Taking the bus: Incorporating public transport timetable data into health care accessibility modelling. *Environment and Planning A*. http://doi.org/10.1068/a4024

MASSACHUSETTS BAY TRANSPORTATION AUTHORITY (MBTA). (2014). Ridership and Service Statistics. Retrieved from http://www.mbta.com/uploadedfiles/documents/2014%20BLUEBOOK%2014th%20Edition.pdf

MAUCH, M., & TAYLOR, B. (1997). Gender, Race, and Travel Behavior: Analysis of Household-Serving Travel and Commuting in San Francisco Bay Area. *Transportation Research Record: Journal of the Transportation Research Board*, *1607*, 147–153. http://doi.org/10.3141/1607-20

MAVOA, S., WITTEN, K., MCCREANOR, T., & O'SULLIVAN, D. (2012). GIS based destination accessibility via public transit and walking in Auckland, New Zealand. *Journal of Transport Geography*, *20*(1), 15–22. http://doi.org/10.1016/j.jtrangeo.2011.10.001

MBTA AND IBI GROUP. (2014). MBTA-Realtime API Documentation. Retrieved from http://realtime.mbta.com/Portal/Content/Documents/MBTA-realtime_APIDocumentation_v2_0_1_2014-09-08.pdf

MESBAH, M., CURRIE, G., LENNON, C., & NORTHCOTT, T. (2012). Spatial and temporal visualization of transit operations performance data at a network level. *Journal of Transport Geography*. http://doi.org/10.1016/j.jtrangeo.2012.07.005

MINOCHA, I., SRIRAJ, P. S., METAXATOS, P., & THAKURIAH, P. (2012). Analysis of Transit Quality of Service and Employment Accessibility for the Greater Chicago, Illinois, Region. *Transportation Research Record Journal of the Transportation Research Board Transportation Research Board of the National Academies*, 20–29. http://doi.org/10.3141/2042-03

MURRAY, A. T., & DAVIS, R. (2001). Equity in Regional Service Provision. *Journal of Regional Science*, *41*(4), 577–600.

NAGATANI, T. (2012). Delay effect on schedule in shuttle bus transportation controlled by capacity. *Physica A: Statistical Mechanics and Its Applications*. http://doi.org/10.1016/j.physa.2012.01.052

PEÑA, C., & MORENO, E. (2014). Delay at Bus Stops of Transmilenio Transport System According to Parameters Measured " in situ " . Case Study Bogotá-Colombia. *Procedia - Social and Behavioral Sciences*, *160*, 121–129. http://doi.org/10.1016/j.sbspro.2014.12.123

PORTER, J. D., KIM, D. S., & GHANBARTEHRANI, S. (2014). *Proof of Concept: Gtfs Data as a Basis for Optimization of Oregon's Regional and Statewide Transit Networks*. Oregon Department of Transportation.

PUBLIC POLICY INSTITUTE OF CALIFORNIA. (2004). *How Much Do Low-Income Californians Pay for Transportation? Research Brief 91*. Public Policy Institute of California. Retrieved from http://www.ppic.org/content/pubs/rb/RB_704LRRB.pdf

PUCHER, J. (1982). Discrimination in Mass Transit. *Journal of the American Planning Association*, *48*(3), 315–326. http://doi.org/10.1080/01944368208976181

PUCHER, J., & RENNE, J. L. (2003). Socioeconomics of Urban Travel: Evidence from the 2001 NHTS. *Transportation Quarterly*, *57*(3), 49–77.

ROOD, T. (1999). *Local Index of Transit Availability (LITA)*.

ROSENBLOOM, S., & ALTSHULER, A. (1977). Equity issues in urban transportation. *Policy Studies Journal*, *6*(1), 29–40.

SABERI, M., ZOCKAIE, A., & FENG, W. (2013). Definition and Properties of Alternative Bus Service Reliability Measures at the Stop Level. *Journal of Public Transportation*, *16*(1).

SANCHEZ, T., & BRENMAN, M. (2008). *Right to Transportation: moving to equity*. American Planning Association.

TAYLOR, B. D., & PAUL, M. O. (1995). Spatial Mismatch or Automobile Mismatch ? An Examination of Race, Residence and Commuting in US Metropolitan Areas. *Urban Studies*, *32*(9), 1453–1473.

THOMOPOULOS, N., & GRANT-MULLER, S. (2013). *Incorporating equity as part of the wider impacts in transport infrastructure assessment: An application of the SUMINI approach*.

TIRACHINI, A. (2013). Estimation of travel time and the benefits of upgrading the fare payment technology in urban bus services. *Transportation Research Part C*, *30*, 239–256. http://doi.org/10.1016/j.trc.2011.11.007

TIRACHINI, A. (2014). The economics and engineering of bus stops: Spacing, design and congestion. *Transportation Research Part A: Policy and Practice*. http://doi.org/10.1016/j.tra.2013.10.010

TRIBBY, C. P., & ZANDBERGEN, P. A. (2012). High-resolution spatio-temporal modeling of public transit accessibility. *Applied Geography*, *34*, 345–355.

WELLS, K., & THILL, J. C. (2012). Do transit-dependent neighborhoods receive inferior bus access? A neighborhood analysis in four U.S. cities. *Journal of Urban Affairs*. http://doi.org/10.1111/j.1467-9906.2011.00575.x

WILSON, W. J. (1990). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

WU, B. M., & HINE, J. P. (2003). A PTAL approach to measuring changes in bus service accessibility. *Transport Policy*. http://doi.org/10.1016/S0967-070X(03)00053-2